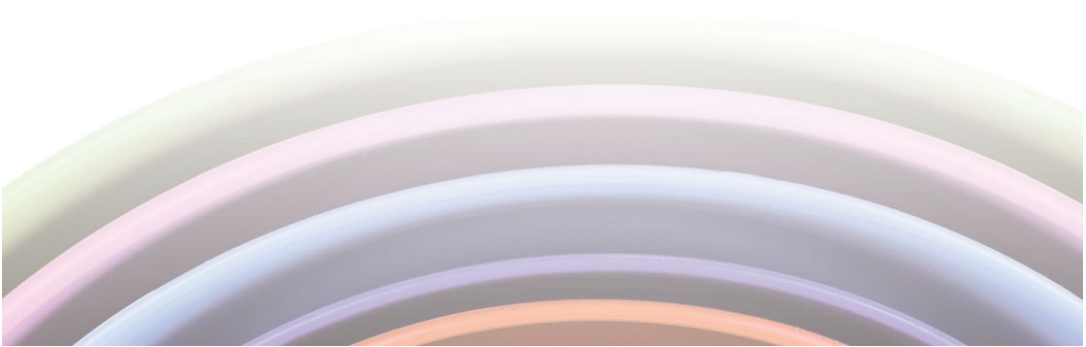


13

Content analysis

Chapter outline

Introduction	289
What are the research questions?	291
Selecting a sample	293
Sampling media	293
Sampling dates	293
What is to be counted?	295
Significant actors	295
Words	295
Subjects and themes	297
Dispositions	298
Coding	298
Coding schedule	298
Coding manual	299
Potential pitfalls in devising coding schemes	303
Advantages of content analysis	304
Disadvantages of content analysis	306
Checklist	307
Key points	308
Questions for review	308





Chapter guide

Content analysis is an approach to the analysis of documents and texts (which may be printed or visual) that seeks to quantify content in terms of predetermined categories and in a systematic and replicable manner. It is a very flexible method that can be applied to a variety of different media. In a sense, it is not a research method in that it is an approach to the analysis of documents and texts rather than a means of generating data. However, it is usually treated as a research method because of its distinctive approach to analysis. This chapter explores:

- the kinds of research question to which content analysis is suited;
- how to approach the sampling of documents to be analysed;
- what kinds of features of documents or texts are counted;
- how to go about *coding*, which is probably the central and most distinctive stage of doing a content analysis;
- the advantages and disadvantages of content analysis.

Introduction

Imagine that you are interested in the amount and nature of the interest shown by the mass media, such as newspapers, in a major news item such as Facebook security or superinjunctions. You might ask such questions as:

- When did news items on this topic first begin to appear?
- Which newspapers were fastest in generating an interest in the topic?
- Which newspapers have shown the greatest interest in the topic?
- At what point did media interest begin to wane?
- Have journalists' stances on the topic changed—for example, in terms of pro- versus anti-Facebook or pro- versus anti-superinjunctions.

If you want to know the answers to research questions such as these, you are likely to need to use content analysis to answer them.

Probably the best-known definition of content analysis is as follows:

Content analysis is a research technique for the objective, systematic and quantitative description of the manifest content of communication. (Berelson 1952: 18)

Another well-known and apparently similar definition is:

Content analysis is any technique for making inferences by objectively and systematically identifying specified characteristics of messages. (Holsti 1969: 14)

It is striking that both these definitions contain a reference to two qualities: objectivity and being systematic. The former quality means that, as with something like an observation schedule (Chapter 12), rules are clearly specified in advance for the assignment of the raw material (such as newspaper stories) to categories. Objectivity in this sense resides in the fact that there is transparency in the procedures for assigning the raw material to categories, so that the analyst's personal biases intrude as little as possible in the process. The content analyst is simply applying the rules in question. The quality of being systematic means that the application of the rules is done in a consistent manner so that bias is again suppressed. As a result of these two qualities, anyone could employ the rules and (hopefully) come up with the same results. The process of analysis is one that means that the results are not an extension of the analyst and his or her personal biases. The rules in question may, of course, reflect the researcher's interests and concerns,

and therefore these might be a product of subjective bias, but the key point is that, once formulated, the rules can be (or should be capable of being) applied without the intrusion of bias.

Berelson's definition also makes reference to 'quantitative description'. Content analysis is firmly rooted in the quantitative research strategy in that the aim is to produce quantitative accounts of the raw material in terms of the categories specified by the rules. The feature of quantification adds to the general sense of the systematic and objective application of neutral rules, so that it becomes possible to say with some certainty and in a systematic way that, for example, during a certain period in which it was potentially at the forefront of media attention, *The Times* and other broadsheet newspapers carried far more coverage of AIDS/HIV than tabloid newspapers, with the interesting exception of the *Sun* (Beharrell 1993).

Two other elements in Berelson's definition are striking especially when juxtaposed against Holsti's. First, Berelson refers to 'manifest content'. This means that content analysis is concerned with uncovering the apparent content of the item in question: what it is clearly about. Holsti makes no such reference, alluding only to 'specified characteristics'. The latter essentially opens the door to conducting an analysis in terms of what we might term 'latent content'—that is, with meanings that lie beneath the superficial indicators of content. Uncovering such latent content means interpreting meanings that lie beneath the surface, such as whether the impression is given that the author construes the AIDS/HIV issue as one solely of concern in relation to the gay community and its sexual practices or as having a broader set of implications that would include heterosexuals. A related distinction is sometimes made between an emphasis on the text (in particular, counting certain words) and an emphasis on themes within the text, which entails searching for certain ideas within the text (Beardsworth 1980).

A second element in Berelson's definition not found in Holsti's is the reference to 'communication'. Berelson's (1952) book was concerned with communication research, a field that has been especially concerned with newspapers, television, and other mass media. Holsti refers somewhat more generally to 'messages', which raises the prospect of a quite wide applicability of content analysis beyond the specific boundaries of the mass media and mass communications. Content analysis becomes applicable to many different forms of unstructured information, such as transcripts of semi- and unstructured interviews (e.g. Bryman, Stephens, and A Campo 1996) and even qualitative case studies of organizations (e.g. Research in focus 13.4). Nor is it necessary for the medium being analysed to be in a printed form. Research has been conducted on:

- the visual images (as well as the text) of women's and men's magazines to examine the degree to which messages about bodily appearance are gendered (Malkin et al. 1999);
- gender roles in animated cartoons (S. N. Davis 2003; see Research in focus 7.7);
- radio and television news programmes (see Research in focus 13.1 for an example);
- speeches, such as the Queen's Speech (John and Jennings 2010) and speeches in the European Parliament (Proksch and Slapin 2010);
- obituaries (Fowler and Bielsa 2007);
- the lyrics of popular songs to reveal changes in the representation of women (Marcic 2000).

However, there is little doubt that the main use of content analysis has been in the examination of printed texts and documents and of mass-media items in particular. In this regard, content analysis is one of a number of approaches to the examination of texts that have been developed over the years (see Key concept 13.1).



Key concept 13.1

What is content analysis?

Content analysis is an approach to the analysis of documents and texts that seeks to quantify content in terms of predetermined categories and in a systematic and replicable manner.

Content analysis can be usefully contrasted with two other approaches to the analysis of the content of communication:

- *Semiotics*. The study/science of **signs**. An approach to the analysis of documents and other phenomena that emphasizes the importance of seeking out the deeper meaning of those phenomena. **Semiotics** is concerned to uncover the processes of meaning production and how signs are designed to have an effect upon actual and prospective consumers of those signs. This approach will be explored in Chapter 23.
- *Ethnographic content analysis*. A term employed by Altheide (1996) to refer to an approach to documents that emphasizes the role of the investigator in the construction of the meaning of and in texts. It is also sometimes referred to as *qualitative content analysis*. As with most approaches that are described as ethnographic, there is an emphasis on allowing categories to emerge out of data and on recognizing the significance for understanding meaning in the context in which an item being analysed (and the categories derived from it) appeared. This approach will be explored in Chapter 23.

When the term 'content analysis' is employed in this chapter, it will be referring to quantitative content analysis—that is, the first of the three forms of analysis referred to in this list, which is the kind of analysis to which Berelson (1952) and Holsti (1969) refer.



What are the research questions?

As with most quantitative research, it is necessary to specify the research questions precisely, as these will guide both the selection of the media to be content analysed and the coding schedule. If the research questions are not clearly articulated, there is a risk that inappropriate media will be analysed or that the coding schedule will miss out key dimensions. Most content analysis is likely to entail several research questions. The research referred to in Research in focus 13.1 was concerned with 'the reporting of social science research in the British mass media'. In itself this is not very specific and hardly directs you to a clear specification of the media to be examined or the development of a coding schedule. However, like most researchers, Fenton et al. (1998) had certain specific research questions in mind, such as:

- How much social science research is reported?
- Do certain mass media report a disproportionate amount of social science research?
- In what locations does social science research tend to get reported (for example, special features rather than general news items)?
- Do some topics receive greater attention than others?
- Are certain social science disciplines favoured by the mass media?
- Do the mass media tend to report research conducted by particular methods?
- What tends to prompt the reporting of social science research?
- Are researchers of a particular status (for example, professors) or from certain institutions (for example, prestigious universities) more likely to receive coverage than others?

Such questions seem to revolve around the questions of: *who* (gets reported); *what* (gets reported); *where* (does the issue get reported); *location* (of coverage within the items analysed); *how much* (gets reported); and *why* (does the issue get reported).

As with much content analysis, the researchers were just as interested in omissions in coverage as in what *does* get reported. For example, details about the status of the researcher(s) and about research methods were frequently omitted. Such omissions are in themselves potentially interesting, as they may reveal what is and is not important to reporters and their editors.

Another kind of issue that is frequently encountered in content analysis is one that was not a concern for the researchers looking at the reporting of social science research:

How far does the amount of coverage of the issue change over time?

This kind of research question is particularly asked by researchers who are keen to note trends in coverage to demonstrate ebbs and flows in media interest. An example of this kind of research is a study by Miller and Reilly (1995) of British newspaper coverage of 'food scares'. They show how newspaper coverage of the salmonella in

eggs scare was the focus of a massive amount of coverage of salmonella over a period of around twenty days in December 1988 following the statements of a British government minister and the ensuing controversy. However, in the eleven months prior to that, salmonella had rarely been in the news, and in the subsequent twelve



Research in focus 13.1

An illustration of content analysis: social science research in the British mass media

Fenton et al. (1998) conducted a study using content analysis of the amount and nature of the reporting of social science research in the British mass media. A sample of eighty-one days of media coverage between 26 May 1994 and 31 March 1995 was taken. The authors state: 'Any media item that mentioned original research conducted by a social scientist or social scientific institution, whether domestic or foreign, was coded, along with any times in which a social scientist (as identified by the item) commented on social issues' (Fenton et al. 1998: 24). The media monitored comprised:

- 12 national newspapers—6 broadsheet and 6 tabloid newspapers; 5 of the 12 newspapers were Sunday papers;
- 4 local newspapers—1 costed and 1 free newspaper from each of Nottingham and Manchester;
- 5 magazines—2 weekly women's publications, 2 monthly women's publications, and 1 monthly men's publication;
- 13 national and local television news programmes covering all four terrestrial television channels broadcasting at the time;
- 6 weekly investigative journalism/social reportage television programmes;
- 5 weekly/daily television talk shows/magazine programmes;
- any prime time ad hoc television programmes deemed relevant;
- national radio news on both Radio 4 and Radio 5 Live;
- 9 national Radio 4 investigative journalism/social reportage/studio talk shows and magazine programmes;
- local radio news in two regions: BBC Radio Nottingham and BBC GMR;
- local radio current affairs on BBC GMR Talkback.

The researchers uncovered 466 cases of research being reported. A further 126 news items were coded in which social scientists acted as pundits. The researchers included the cases of 'punditry' in many of their analyses, because the number of items in which social science research featured in news items was considerably smaller than they had envisaged.

Each news item was coded in terms of a number of features, such as:

- the source of the item (for example, which newspaper or television programme);
- the topic of the research;
- the social science discipline referred to in the news item;
- when no social science discipline was referred to, the 'inferred discipline';
- the professional status of the researcher;
- the main research method employed.

months the amount of coverage was sharply lower than it had been in December 1988. It petered out further over the next four years. However, this decline in coverage occurred in spite of public health evidence that the

incidence of salmonella poisoning was increasing. Similarly, Beharrell (1993) has charted for the period 1988–91 the changes in the amount of UK newspaper coverage of AIDS/HIV issues.



Selecting a sample

There are several phases in the selection of a sample for content analysis. Because it is a method that can be applied to many different kinds of document, the case of applying it to the mass media will be explored here. However, the basic principles have a broader relevance to a wide range of applications of content analysis.

Sampling media

Many studies of the mass media entail the specification of a research problem in the form of ‘the representation of *X* in the mass media’. The *X* may be trade unions, food scares, crime, drink driving, or social science research. But which mass media might one choose to focus upon? Will it be newspapers or television or radio or magazines, or whatever? And, if newspapers, will it be all newspapers or tabloids or broadsheets? And, if both tabloids and broadsheets, will it be all of them and will it include Sunday papers? If it will be a sample of newspapers, including Sunday ones, will these be national or local or both? And will it include free newspapers? And, if newspapers, will all news items be candidates for analysis—for example, would feature articles and letters to the editor be included?

The research reported in Research in focus 13.1 chose to cover a very wide variety of mass media, which is just as well, since the authors were not able to locate a very large number of appropriate items (news items covering social science research). More typically, researchers will opt for one or possibly two of the mass media and may sample within that type or types. Beharrell’s (1993) study of the reporting and representation of AIDS/HIV in the British mass media concentrated on all national and Sunday national newspapers from November 1988 to August 1991, as a result of which a sample of over 4,000 news items was created.

Sampling dates

Sometimes, the decision about dates is more or less dictated by the occurrence of a phenomenon. For example,

Bligh et al. (2004) were keen to explore Weber’s (1947) suggestion that charismatic leadership is most likely to emerge during a period of crisis. They examined the **rhetoric** of President George W. Bush’s speeches before and after the terrorist attacks on the World Trade Center, the Pentagon, and Flight 73 on 11 September 2001. The authors found that not only did his speeches take on a more charismatic rhetoric compared to before the attacks, but that the media portrayal of Bush also tended to incorporate a more charismatic tone. In this case, a key date—9/11—was essentially a given, though there may be an important consideration in deciding at what point the content analysis should cease, since there is an interesting issue concerning the longevity of the charismatic tone of the speeches and media portrayals. The last of the speeches analysed was given on 11 March 2002, which raises an interesting question about how long a charismatic style might be expected to continue following a crisis.

With a research question that entails an ongoing general phenomenon, such as the representation of social science research or crime, the matter of dates is more open. The principles of probability sampling outlined in Chapter 8 can readily be adapted for sampling dates. For example, it is possible to generate a systematic sample of dates by randomly selecting one day of the week and then selecting every *n*th day thereafter. Alternatively, Monday newspapers could provide the first set of newspapers for inclusion, followed by Tuesday the following week, Wednesday the week after, and so on.

One important factor is whether the focus will be on an issue that entails keeping track of representation as it happens, in which case the researcher may begin at any time and the key decision becomes when to stop, or whether it is necessary to go backwards in time to select media from one or more time periods in the past. For example, if Jagger (1998) had wanted to examine whether there had been a marked change in the ways in which men and women represent themselves in dating advertisements (see Research in focus 13.2), she would obviously have needed to examine the columns of earlier



Research in focus 13.2

Finding love

Jagger (1998) reports a content analysis of 1,094 dating advertisements in two Scottish newspapers and two newspapers with a general readership throughout Britain. The sample of advertisements was chosen over two four-week periods: March 1996 and May 1996. Three research questions drove the study.

- What is 'the relative significance of resources and lifestyle choices as identity markers and desirable attributes for men and women' (1998: 799)?
- How far do men and women vary in the degree to which they market themselves and describe their preferred (or ideal) partners in terms of the body?
- How far are 'traditional stereotypes of masculinity and femininity . . . still in operation' (1998: 799)?

Jagger particularly noted the tendency for a considerable percentage of the advertisers to market themselves in terms of their lifestyle choices, a tendency that was not substantially affected by gender. She also found that women were far more likely than men to stress the importance of economic and other resources for their preferred partners. There was also a somewhat greater propensity for women to market themselves in terms of physical appearance. However, men were just as likely to market themselves as 'slim', suggesting that certain norms of bodily shape may no longer be exclusive to one gender. Jagger also found that men frequently market themselves in 'feminine' terms and women in 'masculine' terms—a rather surprising finding. More generally, her results point to the significance of the body in identity construction in modern society for both men and women. In a later publication, Jagger (2005) explores the ways in which consumer culture has had an impact on the representation of the self in this same sample of dating advertisements. She shows that the significance of a person's age and how that is represented varies for men and women.

years. She might have taken comparable samples from ten and twenty years earlier and perhaps even beyond. Warde (1997) was interested in changes in the representation of food (what should be eaten and how it should be eaten) in the food columns of women's magazines. He writes:

My primary sources were the five most widely read women's weekly magazines and the five most widely read monthly magazines in each of two twelve-month periods in 1967–8 and 1991–2. The magazines were sampled at the mid-point of the months of November, February, May and August in each year, in order to control for seasonal variation in the contents of food columns. . . . From the selected magazines I drew a systematic sample of recipes. This produced 114 recipes in the earlier year, 124 in the later period[,] which, given their random selection, should be sufficient to make some generalizations about recipes and any changes over time. (Warde 1997: 44–5)

Three points stand out in this passage. First, there is the concern, which has just been alluded to, with being able to establish change by tracking back in time to earlier issues of the mass medium being analysed. Second, Warde wanted to ensure that the magazines were selected from four different points in each of the two twelve-month periods in order to ensure that seasonal factors did not overly influence the findings. If he had selected magazines just from November, there might have been a preoccupation with Christmas fare, while findings from magazines from a summer month might have been affected by the greater availability of certain foods, such as particular fruit. A decision was made to cover the four seasons in order to reduce the impact of such factors. Third, there is a clear concern to enhance the representativeness of the recipes and therefore the generalizability of the findings by using a probability sampling method in the form of a systematic sample (see Chapter 8).



What is to be counted?

Obviously, decisions about what should be counted in the course of a content analysis are bound to be profoundly affected by the nature of the research questions under consideration. Content analysis offers the prospect of different kinds of 'units of analysis' being considered. The following kinds of units of analysis are frequently encountered and can be used as guides to the kinds of objects that might be the focus of attention. However, what you would actually *want* or *need* to count will be significantly dictated by your research questions.

Significant actors

Particularly in the context of mass-media news reporting, the main figures in any news item and their characteristics are often important items to code. These considerations are likely to result in such persons as the following being recorded in the course of a content analysis.

- What kind of person has produced the item (for example, general or specialist news reporter)?
- Who is or are the main focus of the item (for example, politician, expert, government spokesperson, or representative of an organization)?
- Who provides alternative voices (for example, politician, expert, government spokesperson, representative of an organization, or person in the street)?
- What was the context for the item (for example, interview, release of a report, or an event such as an outbreak of hostilities or a minister's visit to a hospital)?

In the case of the content analysis of the reporting of social science research in the mass media (see Research in focus 13.1), the significant actors included:

- the author of the item (for example, type of correspondent);
- the type of item (for example, in the case of the press, whether the research was reported in a general article, feature article, or some other context);
- the details of the researcher who was most prominent in the item (for example, personal details, status, and whether he or she was acting as a researcher or pundit in the context of the item);
- what prompted the item (for example, launch of a report, new research initiative, or a conference);

- the details of the main (if any) commentators on the research;
- any other actors.

The chief objective in recording such details is to map the main protagonists in news reporting in an area and to begin to reveal some of the mechanics involved in the production of information for public consumption.

Words

While it may seem a dull activity, the counting of the frequency with which certain words occur is sometimes undertaken in content analysis. However, the use of some words rather than others can often be of some significance, because it can reveal the predilection for sensationalizing certain events. For example, Dunning et al. (1988) have noted a tendency for the British press in the late 1960s to sensationalize the reporting of disturbances at football matches. The use of such emotive words as 'hooligan' and 'lout', along with inferences about 'war', are clearly significant. Alternative, less dramatic terms could have been used. Instead, the florid accounts of violence at football matches may have created the kind of 'moral panic' that Cohen (1973) has written about. While Dunning et al. did not conduct a content analysis of such news reporting, their analysis points to the potential significance of the choice of certain words and the neglect of others. Jagger's (1998) study of dating advertisements (see Research in focus 13.2) counted such words as 'slim' and 'non-smoker' to uncover some of the characteristics deemed desirable.

A variation on the search for the occurrence of certain words is the examination of keywords in the context of other words. Hansen (1995) provides an interesting example in a study of the reporting of BSE (mad cow disease) in the *Today* newspaper. He found that news items in which BSE appeared often included other prominent words, such as 'beef', 'food', 'meat', 'pounds [£]', 'Government', and 'ban'. The frequency with which such words accompanied 'mad', 'cow', and 'disease' led Hansen (1995: 159) to infer that 'there is a clear indication . . . that the coverage focuses on: (1) the danger which the disease poses for a key component of the British diet, (2) on costs, and (3) on what the government is doing or should be doing about it'. The examination of such

key accompanying words can then be a springboard for a more thematic analysis (see the section on 'Subjects and themes' below). Hansen notes that the frequent occurrence of certain words suggests certain types of 'discourse'. For example, he identifies a discourse concerned with the *transmission* of BSE (such words as: infected, catch,

spread, etc.) and a discourse emphasizing risk and threat (such words as: safe, risk, fears, danger, etc.). Tips and skills 'Counting words in electronic news reports' discusses the use of this kind of analysis in the context of electronic newspapers. Some examples of content analysis that emphasize words can be found in Research in focus 13.3.



Tips and skills

Counting words in electronic news reports

The growing availability of the printed news media in electronic form, such as CD-ROM and electronic databases, greatly facilitates the search for and counting of keywords in this kind of context. Most of the main UK newspapers and many overseas ones are available in electronic format either through their own websites or through a website like British Media Online:

www.wrx.zen.co.uk/britnews.htm (accessed 7 October 2010),

which acts as a launch pad for a host of different electronic newspapers. The newspapers can then usually be searched for keywords and phrases. You will probably need a password (for example, Athens authentication) to access them, especially if you are seeking to access them away from your university or college.

Further, some simple analyses can be conducted using LexisNexis. This database comprises newspaper articles from a wide variety of newspapers. For example, Beardsworth and Bryman (2004) used Lexis Nexis Professional to search this database for the years 1985 to 2002 for the incidence of keywords relating to BSE, such as: Bovine Spongiform Encephalopathy; BSE; CJD; and Creutzfeld Jacob Disease. They chart the number of reports using at least one of these terms in each newspaper per annum. Their analysis shows how there was a small surge of media reporting around 1990 and then a huge 'spike' in the mid-1990s, when it attracted a great deal of media interest, and then a trailing-off in the incidence of reporting. They also show how the statistics for beef and veal consumption in the UK display a corresponding pattern. This article demonstrates the use of a simple counting procedure that can be quite revealing and that is greatly facilitated by the availability of electronic (and in this case online) mass media. However, caution is needed in the use of LexisNexis. Deacon (2007) has drawn attention to several drawbacks in its use relative to conventional manual searching. For example, he notes that there is considerable loss of information because the visual elements of media reporting are not included and because news articles are treated in isolation and not in relation to other news articles.



Research in focus 13.3

Computerized keyword analysis

Seale et al. (2006) report the results of a study in which he retrieved all postings on a single day (20 April 2005) to online forums of the two most prevalent websites in the UK concerned with breast and prostate cancer. Keywords in this research were words that occur with unusual frequency compared with other words in the corpus that is analysed. In this case, Seale and his colleagues were interested in keyword frequency in the breast cancer postings as compared to the prostate cancer postings. The search for keywords was undertaken with specialist software called *Wordsmith*:

www.lexically.net/wordsmith/ (accessed 11 October 2010).

Seale et al. (2006: 2582) then 'used this quantitative information to facilitate an interpretive, qualitative analysis focusing on the meanings of word clusters associated with keywords'. Thus, Seale and his colleagues use this quantitative analysis of words as a springboard for a more probing qualitative examination of the links between the words. They found that men with prostate cancer were more likely to use words associated with research (for example, 'study'), treatment (for example, 'radical prostatectomy'), and tests and diagnosis (for example, 'biopsy') compared to women discussing breast cancer. By contrast, women discussing breast cancer were more likely to use keywords associated with feelings (for example, 'scared') and people (for example, 'family') and also to use 'superlatives' (for example, 'amazing').

In a later study, the same analytic approach was applied to postings in April 2005 to Internet forums concerning breast cancer, prostate cancer, and sexual health (Seale et al. 2010). The analysis was compared to a parallel examination of qualitative interviews with people discussing their experiences of breast and prostate cancer and their sexual health concerns (a total of 140 interviews). The authors report that what they call the 'orientation' of the two formats differed. In interviews, there was an emphasis on the reporting of the past, presumably because the interview format encouraged narratives about the past experience of the illness. This is a focus that has been a particular feature of studies of health and illness that have been influenced by **narrative analysis** and the notion of 'illness narratives' (see Chapter 24). The Internet postings tended to be about the exchange of information and support and therefore dealt more directly with the immediate experience of illness. Seale et al. propose that this means that the Internet postings provide superior access to the experience of illness as the accounts are not produced for the benefit of an interviewer. The Internet postings also come across as more spontaneous and as allowing greater freedom for issues to be raised that relate to the sufferer's experience (for example, the frequent discussion of chocolate among breast cancer postings), whereas the interviewer's influence was very evident (for example, 'relationship' or 'relationships' were used in questions 622 times). The authors argue that the tendency for interviewers to seek to elicit narratives about the illness and with particular issues in mind tends to militate against direct expressions of illness experiences, although the ability to probe, which can yield significant data, is something missing in the postings. Seale et al. argue that, while they are not flawless as sources of data about illness, particularly for research into sensitive topics where there are concerns about the truthfulness of reports in interviews, Internet postings about illness experiences warrant serious consideration, especially when the time and cost involved in sampling, conducting, and transcribing qualitative interviews are borne in mind.

Subjects and themes

Frequently in a content analysis the researcher will want to code text in terms of certain subjects and themes. Essentially, what is being sought is a categorization of the phenomenon or phenomena of interest. For example, in the case of the content analysis of the reporting of social science research in the British mass media, Fenton et al. (1998) were concerned to classify the main social science discipline that formed the backcloth to the research being reported (Research in focus 13.1). This entailed a classification into one of seven types: sociology; social policy; economics; psychology; business and management; political science; and interdisciplinary. Research drawing on other social science disciplines was not included in the study. Another topic was the methodology of the research being reported, which resulted in research being classified in terms of such categories as: survey/mail

questionnaire; interview; ethnography; and, of course, content analysis.

However, while such categorizations are often relatively straightforward, when the process of coding is thematic, a more interpretative approach needs to be taken. At this point, the analyst is searching not just for manifest content but for latent content as well. It becomes necessary to probe beneath the surface in order to ask deeper questions about what is happening. In the research on social science research in the mass media (see Research in focus 13.1), each reported study was classified in terms of the subject area of the research (Fenton et al. 1998). Sixty-two categories were employed and were grouped into seven main areas: UK and overseas economy; UK and overseas government politics and policy; social integration and control; health; demographics; social analysis—general; and lifestyles. Another example of having to interpret derives from the researchers'

examination of what they refer to as the 'inferred discipline' involved in a reported piece of research. This issue arose because, in many cases (nearly 60 per cent of all news items in which social science research was reported), no social science discipline was mentioned in the news report, so that it was necessary to infer which discipline was involved.

Dispositions

A further level of interpretation is likely to be entailed when the researcher seeks to demonstrate a disposition in the texts being analysed. For example, it may be that the researcher wants to establish whether journalists, in the reporting of an issue in the news media, are favourably inclined or hostile towards an aspect of it, such as their stances on the government's handling of a food scare crisis. In the case of the study by Fenton et al. (1998; see Research in focus 13.1) of the reporting of social science research, each item was coded in terms of

whether the editorial commentary on the research was positive, negative, or merely descriptive. In many cases, it was necessary to infer whether the editorial commentary was implicitly positive or negative if there were no manifest indications of such value positions. Such an analysis entails establishing whether a judgemental stance can be discerned in the items being coded and the nature of the judgement.

Another way in which dispositions may be revealed in content analysis is through the coding of ideologies, beliefs, or principles. Jagger (1998) coded dating advertisements in terms of whether gendered stereotypical categories of masculinity and femininity were employed when advertisers were describing themselves (see Research in focus 13.2). She came up with the surprising finding that women were more likely than men to advertise themselves in terms of a masculine stereotype, while men were more likely to advertise themselves in terms of a feminine stereotype.



Coding

As much of the foregoing discussion has implied, coding is a crucial stage in the process of doing a content analysis. There are two main elements to a content analysis coding scheme: designing a coding schedule and designing a coding manual. To illustrate its use, imagine a student interested in crime reporting in a local newspaper. The student chooses to focus on the reporting of crimes subject to court proceedings and where the victim is a person rather than an organization. To simplify the issue we will just have the following variables:

1. nature of the offence;
2. gender of perpetrator;
3. social class of perpetrator;
4. age of perpetrator;
5. gender of victim;
6. social class of victim;
7. age of victim;
8. depiction of victim;
9. position of news item.

Content analysts would normally be interested in a much larger number of variables than this, but a simple illustration like this can be helpful for getting across the

operation of a coding schedule and a coding manual. Also, it is quite likely that the student would want to record the item so that the details of more than one offender and more than one victim are included. In other words, very often a crime will entail multiple perpetrators and/or victims, so that the details of each of the key figures (age, gender, occupation, depiction of victim) would need to be recorded. However, to keep the illustration simple, just one perpetrator and victim is assumed.

Coding schedule

The coding schedule is a form onto which all the data relating to an item being coded will be entered. Figure 13.1 provides an example of a coding schedule. The schedule is very much a simplification in order to facilitate the discussion of the principles of coding in content analysis and of the construction of a coding schedule in particular.

Each of the columns in Figure 13.1 is a dimension that is being coded. The column headings indicate the dimension to be coded. The blank cells on the coding form are the places in which codes are written in. One form would be used for each media item that was coded. The codes can then be transferred to a computer data file for analysis with a software package like SPSS (see Chapter 16).

Figure 13.1

Coding schedule

Case number	Day	Month	Year	Nature of offence I	Gender of perpetrator	Occupation of perpetrator	Age of perpetrator	Gender of victim	Occupation of victim	Age of victim	Depiction of victim	Nature of offence II	Position of news item

Coding manual

On the face of it, the coding schedule in Figure 13.1 seems very bare and does not appear to provide much information about what is to be done or where. This is where the coding manual comes in. The coding manual is a statement of instructions to coders that also includes all the possible categories for each dimension being coded. It provides: a list of all the dimensions; the different categories subsumed under each dimension; the numbers (that is, *codes*) that correspond to each category; and guidance on what each dimension is concerned with and any factors that should be taken into account in deciding how to allocate any particular code to each dimension. Figure 13.2 provides the coding manual that might correspond to the coding schedule in Figure 13.1. A coding manual includes all the dimensions that would be employed in the coding process, indications of the guidance for coders, and the kinds of lists of categories that were created for each dimension.

The coding manual includes the occupation of both the perpetrator and the victim. It uses Goldthorpe's social class categorization and is based on the summary by Marshall et al. (1988: 22). To this scheme have been added three further categories that might be used in newspapers: unemployed; retired; and housewife. There is also a category of 'other'. The offences are categorized in terms of those used by the police in recording crimes notified to them according to Home Office rules. Much finer distinctions could be used, but, since the student may not be planning to examine a large sample of news items, broader categories might be preferable. They have the further advantage of being comparable to the Home Office data. Recorded crime statistics have been criticized for their lack of reliability and validity (see Chapter 14), but the comparison between such data

and the reporting of crime in local newspapers would be potentially illuminating.

The coding schedule and manual permit two offences to be recorded when an incident entails more than one offence. If there are more than two, the student has to make a judgement concerning the most significant offence. The student should also treat as the first offence the main one mentioned in the article.

The coding manual is crucial because it provides coders with complete listings of all categories for each dimension they are coding and guidance about how to interpret the dimensions. It is on the basis of these lists and guidance that a coding schedule of the kind presented in Figure 13.1 will be completed. Even if you are a lone researcher, such as a student conducting a content analysis for a dissertation or thesis, it is important to spend a lot of time providing yourself with instructions about how to code. While you may not face the problem of **inter-coder reliability**, the issue of **intra-coder reliability** is still significant for you (see below).

Plates 13.1 and 13.2 provide examples of the kind of article that might appear. Both are from the *Nottingham Evening Post*.

The coding of the incidents would then be added to coding schedules, as shown in Figure 13.1, and the data from each form would then be entered as a row of data in a computer program like SPSS.

The coding of the incident in Plate 13.1 would appear as in Figure 13.3; the data would appear as follows:

123 27 12 02 1 1 17 46 1 17 -1 3 16 2

Note that the news item in Plate 13.2 contains a second offence, which has been coded as 16 under 'Nature of offence II'. Figure 13.4 contains the form that would be completed for the item in Plate 13.2. The following row of data would be created:

Figure 13.2

Coding manual

Nature of offence I

1. Violence against the person
2. Sexual offences
3. Robbery
4. Burglary in a dwelling
5. Burglary other than in a dwelling
6. Theft from a person
7. Theft of pedal cycle
8. Theft from shops
9. Theft from vehicle
10. Theft of motor vehicle
11. Vehicle interference and tampering
12. Other theft and handling stolen goods
13. Fraud and forgery
14. Criminal damage
15. Drug offences
16. Other notifiable offences

Gender of perpetrator

1. Male
2. Female
3. Unknown

Occupation of perpetrator

1. I Higher grade professionals, administrators, and officials; managers in large establishments; large proprietors
2. II Lower-grade professionals, administrators, and officials; higher-grade technicians; managers in small business and industrial establishments; supervisors of nonmanual employees
3. IIIa Routine nonmanual employees in administration and commerce
4. IIIb Personal service workers
5. IVa Small proprietors, artisans, etc., with employees
6. IVb Small proprietors, artisans, etc., without employees
7. IVc Farmers and smallholders; self-employed fishermen
8. V Lower-grade technicians, supervisors of manual workers
9. VI Skilled manual workers
10. VIIa Semi-skilled and unskilled manual workers (not in agriculture)

11. VIIb Agricultural workers
12. Unemployed
13. Retired
14. Housewife
15. Student
16. Other
17. Unknown

Age of perpetrator

Record age (–1 if unknown)

Gender of victim

1. Male
2. Female
3. Unknown
4. Organization (if victim is an organization as in fraud cases)

Occupation of victim

Same as for occupation of perpetrator
If not applicable, code as 99

Age of victim

Record age (–1 if unknown; –2 if not applicable)

Depiction of victim

1. Victim responsible for crime
2. Victim partly responsible for crime
3. Victim not at all responsible for crime
4. Not applicable

Nature of offence II (code if second offence mentioned in relation to the same incident; code 0 if no second offence)
Same as for Nature of offence I

Position of news item

1. Front page
2. Inside
3. Back page

Plate 13.1

Reporting a crime in local newspapers I

Dogs fighting led to assault

Owner guilty of chain attack and cruelty

By SEAN KENNY

IT began with a fight between two dogs — but ended when one owner faced assault and animal cruelty charges.

Conrad Aaron Martin, 46, of Leybourne Drive, Bestwood, attacked a dog and its owner with a chain when they met while out for a morning walk, Nottingham magistrates heard.

He was out with his English and Staffordshire bull terrier cross, on May 27 on the Hucknall Embankment in Bestwood when his pet met Rocco Lamagna's German shepherd dog, Lupo, out for his morning walk.

The terrier locked his teeth around Lupo's front leg and Mr Lamagna, of Bestwood, asked Martin to control his dog.

But, the court was told, Martin replied: "I'm going to kill you... and your dog."

Mr Lamagna told the court: "He got a massive chain and hit me and the dog with it."

"I tried to bend over my dog, I couldn't do anything. My dog was bleeding from his eye."

Mr Lamagna told the court he had not taken his dog along that route since the incident.

Talking to the *Post* after the trial, Mr Lamagna said: "The attack was totally unexpected."

"I'd just come out of hospital and was taking Lupo for a walk when his dog went for mine."

Mr Lamagna controlled his dog and tried to protect him: "Lupo's eye was bleeding and I tried to cover him to stop him getting bitten again."

"The vet's bill cost me £65."

"I used to walk Lupo down there all the time, but I never go there now."

Denied

The court heard that after Mr Lamagna reported the attack to the police, investigations led to an identity parade being organised.

Mr Lamagna immediately picked out Martin, the court was told.

Martin had denied the charges of



FIGHT VICTIMS: Rocco Lamagna and his dog Lupo.

C241202S07-1

assault and cruelty to animals, but was found guilty.

He told the court: "There's nothing I wouldn't do to defend my dog."

Magistrate Mrs Ann Allison said: "We accept Mr Lamagna was certain when he picked you out at the identity

parade."

Martin has two previous convictions for affray whilst walking his dog, in November and December last year.

His sentence will be decided by the crown court.

Plate 13.2

Reporting a crime in local newspapers II

Landlords hit out as pub attacker avoids jail



PUB ASSAULT: Injured landlord Paul Buxton and, left, his assailant Tom Turton

ANGRY licensees have blasted the "deplorable" sentence given to a violent customer.

Tom Turton 24, poked his thumbs into landlord Paul Buxton's eyes after being asked to leave the Lady Bay pub in Trent Boulevard.

Turton, of Fourth Avenue, Sherwood Rise, was ordered to do 260 hours community service by Nottingham magistrates after admitting common assault. He was also told to pay £170 compensation and £55 costs.

The sentence has outraged members of the West Bridgford Pubwatch scheme, of which Mr Buxton is a member.

Letter

Today they were sending a letter to the Chairman of the Magistrates Bench which reads: "We are writing to express our disgust at the leniency of the sentence given to the accused Tom Turton. What kind of message are we giving out here?"

The landlords described the landing of the case, including numerous adjournments, as "deplorable". The Pubwatchgroup is also taking legal advice about whether to appeal against the decision.

Alan Merryweather, the Pubwatch vice-chairman who runs the Tet Match Inn in West Bridgford, said: "I don't know what guidelines the magistrates have but Turton should have got a six month custodial sentence. We feel that reflects the severity of the injuries."

Assaulted landlord Paul Buxton said: "I don't think his sentence was stiff enough."

"He needs a short, sharp shock — but the message that's gone out is smack a landlord and you'll get away with it."

"No-one should go to work and get punched or poked in the eyes."

Magistrates heard that on Sunday, March 9, Turton had spent all afternoon and evening drinking in the Lady Bay pub, downing six or seven pints of lager.

At closing time, he became aroused when Mr Buxton jokingly told him to leave.

Dennis Quinn, prosecuting, told the court: "He deliberately stuck both thumbs in each of Mr Buxton's eyes."

Mr Buxton said he could feel the thumbs in his eye sockets.

The landlord suffered cuts to the rear of his eyes. It has left him with occasional blurred vision and needing eye drops, but there will be no lasting damage.

Mr Turton was the first person to be banned from all 22 pubs within the West Bridgford Pubwatch.

Sundeep Sorin, in mitigation, said it was Turton's first violent offence and he was very remorseful.

Disgust at 'lenient' sentence



ANGRY: Mr Buxton with the letter sent to magistrates and, above, the Lady Bay pub

Graham Hooper, Justice Clerk at Nottingham Magistrates' Court, said the maximum penalty for common assault is a £5,000 fine or six months in prison.

He added: "When a licensee or someone going about their business serving the public is assaulted that's seen as a fact which makes the offence more serious."

"But magistrates are required to look at the circumstances of the offender and consider each case on its merits."

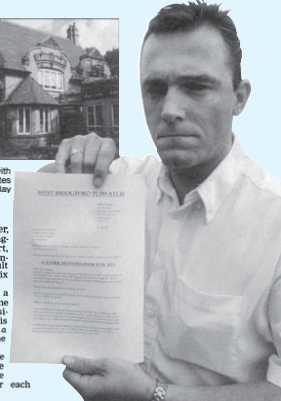


Figure 13.3

Completed coding schedule for news item in Plate 13.1

Case number	Day	Month	Year	Nature of offence I	Gender of perpetrator	Occupation of perpetrator	Age of perpetrator	Gender of victim	Occupation of victim	Age of victim	Depiction of victim	Nature of offence II	Position of news item
123	27	12	02	1	1	17	46	1	17	-1	3	16	2

Figure 13.4

Completed coding schedule for news item in Plate 13.2

Case number	Day	Month	Year	Nature of offence I	Gender of perpetrator	Occupation of perpetrator	Age of perpetrator	Gender of victim	Occupation of victim	Age of victim	Depiction of victim	Nature of offence II	Position of news item
301	04	07	03	1	1	17	24	1	5	-1	3	0	2

301 04 07 03 1 1 17 24 1 5 -1 3 0 2

Forms like these would be completed for each news item within the chosen period or periods of study.

Potential pitfalls in devising coding schemes

There are several potential dangers in devising a content analysis coding scheme, and they are very similar to the kinds of consideration that are involved in the design of structured interview and structured observation schedules.

- *Discrete dimensions.* Make sure that your dimensions are entirely separate; in other words, there should be no conceptual or empirical overlap between them. For example, in the research presented in Research in focus 13.1, it is necessary to be clear about the difference between discipline and topic of the research, even though they have a similar ring.
- *Mutually exclusive categories.* Make sure that there is no overlap in the categories supplied for each dimension. If the categories are not mutually exclusive, coders will be unsure about how to code each item.
- *Exhaustive categories.* For each dimension, all possible categories should be available to coders.
- *Clear instructions.* Coders should be clear about how to interpret what each dimension is about and what factors to take into account when assigning codes to each category. Sometimes, these will have to be very elaborate. Coders should have little or no discretion in how to allocate codes to units of analysis.
- *Clarity about the unit of analysis.* For example, in the study of social science research in the mass media (see Research in focus 13.1), the authors found it necessary to operate with a clear distinction between the media item (e.g. a newspaper article) and the social science research being reported. Thus some of the data recorded were to do with the media item; other data recorded were to do with the research being referred to. The researchers' content analysed up to three social science research profiles per media item, because often the media reporting of research referred to more than one investigation. In the study reported in Research in focus 7.7, 167 cartoons were sampled, but it was the *characters* in those cartoons that provided the units of analysis. As a result, characteristics of the 478 characters were coded and analysed.

Similarly, in the imaginary study of the media reporting of crime in the local press, more than one offence per media item could be recorded. You need to be clear about the distinction between the media item (e.g. a newspaper article) and the topic being coded (an offence). In practice, a researcher is interested in both but needs to keep the distinction in mind.

In order to be able to enhance the quality of a coding scheme, it is highly advisable to pilot early versions of the scheme. Piloting will help to identify difficulties in applying the coding scheme, such as uncertainty about which category to employ when considering a certain dimension or discovering that no code is available to cover a particular case. Piloting will also help to identify

any evidence that one category of a dimension tends to subsume an extremely large percentage of items. If this occurs, it may be necessary to consider breaking that category down so that it allows greater discrimination between the items being analysed.

The reliability of coding is a further potential area of concern. Coding must be done in a consistent manner. As with structured observation, coding must be consistent between coders (*inter-coder reliability*), and each coder must be consistent over time (*intra-coder reliability*). An important part of piloting the coding scheme will be testing for consistency between coders and, if time permits, intra-coder reliability. The process of gauging reliability is more or less identical to that briefly covered in the context of structured observation in Key concept 12.3.



Advantages of content analysis

Content analysis has several advantages, which are outlined below.

- Content analysis is a very transparent research method. The coding scheme and the sampling procedures can be clearly set out so that replications and follow-up studies are feasible. It is this transparency that often causes content analysis to be referred to as an objective method of analysis.
- It can allow a certain amount of longitudinal analysis with relative ease. Several of the studies referred to above allow the researcher to track changes in frequency over time (Warde 1997; Bligh et al. 2004). The time periods can have long spans, such as Warde's (1997) analysis of women's magazines over two time periods over twenty years apart, or quite short time spans, such as the research by Bligh et al. (2004) on President Bush's speeches in the months before and after 9/11. Similarly, changes in emphasis in crime reporting in newspapers over two very different periods can be examined.
- Content analysis is often referred to favourably as an *unobtrusive method*, a term devised by Webb et al. (1966) to refer to a method that does not entail participants in a study having to take the researcher into account (see Key concept 14.3). It is therefore a *non-reactive method* (see Key concept 12.4 and 14.3). However, this point has to be treated with a little caution. It is certainly the case that, when the focus of a content analysis is upon things such as newspaper articles or television programmes, there is no reactive effect. Newspaper articles are obviously not written in the knowledge that a content analysis may one day be carried out on them. On the other hand, if the content analysis is being conducted on documents such as interview transcripts or ethnographies (e.g. Hodson 1996; see Research in focus 13.4), while the process of content analysis does not itself introduce a reactive effect, the documents may have been at least partly influenced by such an effect.
- The unobtrusiveness of content analysis can be significant for many students. For students conducting research for an undergraduate or postgraduate project (perhaps for a dissertation), content analysis has the advantage that it does not usually require them to undergo the same level of ethical scrutiny that is common for students selecting methods that require research participants. This is not to suggest that students should select research methods on the basis that they allow them to avoid increasingly difficult ethical surveillance regimes, particularly given the arguments in Chapters 1 and 4 about the need to tailor research methods to research questions. However, it does point to a consideration that is worth bearing in mind when thinking about how to approach a student research project.
- It is a highly flexible method. It can be applied to a wide variety of different kinds of unstructured textual

information. While content analysis is often associated with the analysis of mass-media outputs, it has a much broader applicability than this. Research in focus 13.4 presents an illustration of a rather unusual but none the less interesting application of content analysis.

- Content analysis can allow information to be generated about social groups to which it is difficult to gain access. For example, most of our knowledge of the social backgrounds of elite groups, such as senior clergy, company directors, and top military personnel, derives from content analyses of such publications as *Who's Who* and *Burke's Peerage* (Bryman 1974).



Student experience

The significance of the transparency of content analysis

For her research on the front covers of women's magazines, Sarah Hanson felt that the transparency of the research method she had chosen was significant. She writes:

My supervisor, Kristin, influenced my research methods a lot, as, having never done a study quite so large before, I needed a good way to organize my research and writing. By introducing Bryman's *Social Research Methods* (Second Edition) and the chapter on Content Analysis into my dissertation, I was able to do a lot of research within a controlled environment; my total word count for the research in my appendix tables was over 2,000 words. Content analysis allows for many different types of research and data collection to be easily carried out and studied, allowing you to work methodically throughout the decoding and analysis. By including content analysis in my dissertation, it allowed for easy reference throughout the text and allowed the marker to see the 'mathematical' workings, providing proof of how I came to my final conclusions.

It is nice to know one's work has an impact!



To read more about Sarah's research experiences, go to the Online Resource Centre that accompanies this book at: www.oxfordtextbooks.co.uk/orc/brymansrm4e/



Research in focus 13.4

A content analysis of qualitative research on the workplace

Hodson (1996: 724) reports the results of a content analysis of 'book-length ethnographic studies based on sustained periods of direct observation'. There is an excellent website in connection with the Workplace Ethnography Project, which can be found at:

www.sociology.ohio-state.edu/rdh/Workplace-Ethnography-Project.html (accessed 7 October 2010).

As a method, ethnography, which will be explored in detail in Chapter 19, entails a long period of participant observation in order to understand the culture of a social group. Hodson's content analysis concentrated on ethnographic studies of workplaces that had been published in book form (that is, published articles were excluded because they rarely included sufficient detail). Thousands of case studies were assessed for possible inclusion in the sample. The sample was made up of studies from different countries. According to the Workplace Ethnography Project website:

www.sociology.ohio-state.edu/rdh/Workplace-Ethnography-Project.html (accessed 7 October 2010),

the current and probably final tally is described as follows: 'The study generated 204 ethnographic cases. These cases were derived from 156 separate books since the observations reported in some books allowed the coding of multiple cases.'

In Hodson (1996) each case was coded in terms of one of five types of workplace organization (craft, direct supervision, assembly line, bureaucratic, and worker participation). This was the independent variable. Various dependent variables and 'control' variables (variables deemed to have an impact on the relationships between independent and dependent variables) were also coded.

Here are two of the variables and their codes.

WORKERS:

Job satisfaction

1 = very low; 2 = moderately low; 3 = average; 4 = high; 5 = very high

Autonomy

1 = none (the workers' tasks are completely determined by others, by machinery or by organizational rules); 2 = little (workers occasionally have the chance to select among procedures or priorities); 3 = average (regular opportunities to select procedures or set priorities within definite limits); 4 = high (significant latitude in determining procedures and setting priorities); 5 = very high (significant interpretation is needed to reach broadly specified goals) (Hodson 1996: 728)

Hodson's findings suggest that some pessimistic accounts of worker participation schemes (for example, that they do not genuinely permit participation and do not necessarily have a beneficial impact on the worker) are incomplete. A more detailed treatment of this research can be found in Hodson (1999). Since this early 1996 publication, many others have been published in major journals, including the article on social fulfilment at the workplace (Hodson 2004), which was referred to in Chapter 4.

Not only does the website provide a list of publications deriving from the project (including downloadable pdf files of most of the articles) and all the coding information; you can also download the data into SPSS, which is the software that will be covered in Chapter 16.



Disadvantages of content analysis

Like all research techniques, content analysis suffers from certain limitations, which are described below.

- A content analysis can only be as good as the documents on which the practitioner works. Scott (1990) recommends assessing documents in terms of such criteria as: authenticity (that the document is what it purports to be); credibility (whether there are grounds for thinking that the contents of the document have been or are distorted in some way); and representativeness (whether the documents examined are representative of all possible relevant documents, as, if certain kinds of document are unavailable or no longer exist, generalizability will be jeopardized). These kinds of consideration will be especially important to bear in mind when a content analysis is
- being conducted on documents like letters. These issues will be explored in further detail in Chapter 23.
- It is almost impossible to devise coding manuals that do not entail some interpretation on the part of coders. Coders must draw upon their everyday knowledge as participants in a common culture in order to be able to code the material with which they are confronted (Cicourel 1964; Garfinkel 1967). To the extent that this occurs, it is questionable whether it is justifiable to assume a correspondence of interpretation between the persons responsible for producing the documents being analysed and the coders (Beardsworth 1980).
- Particular problems are likely to arise when the aim is to impute latent rather than manifest content. In

searching for traditional markers of masculinity and femininity (Research in focus 13.2) or inferring a social science discipline (Research in focus 13.1), the potential for an invalid conjecture being made is magnified.

- It is difficult to ascertain the answers to 'why?' questions through content analysis. For example, Jagger (1998: 807) found that 'the body of their partner, its attractiveness, shape and size, was of less importance to advertisers when in the buying mode [advertising for a partner]' than when in the selling mode (marketing oneself) (see Research in focus 13.2). Why? And finding that this was true of both men and women is, as Jagger suggests, even more surprising. But again, why? We can speculate, but our suggested answers can usually only be speculations. Similarly, Fenton et al. (1998) found that sociology was only the fourth most common discipline to be referred to when social science research was being reported (see Research in focus 13.1). However, Fenton et al. also inferred subject disciplines when they were not referred to explicitly by journalists. Sociology was by far the most frequent *inferred* discipline. Again, while this is an interesting finding, the reasons for it can only be speculated about (Fenton et al. 1998). Sometimes, users of content analysis have been able to shed some light on 'why?' questions raised by their investigations by conducting additional data-collection exercises. Such exercises might include qualitative content analysis (e.g. Glasgow University Media Group 1976) and/or interviews with journalists and others (e.g. Fenton et al. 1998).
- Content analytic studies are sometimes accused of being atheoretical. It is easy to see why an atheoretical approach might arise. The emphasis in content analysis on measurement can easily and unwittingly result in an accent being placed on what is measurable rather than on what is theoretically significant or important. However, content analysis is not necessarily atheoretical. Jagger (1998, 2005) placed her findings on dating advertisements in the context of current ideas about consumerism and the body. Fenton et al. (1998) conducted their content analysis within an overall approach that stressed the importance of studying the mass communication process from inception to reception and the importance of power and contestation within that process. Hodson's (1996) content analysis of workplace ethnographies was underpinned by theoretical ideas deriving from the work of influential writers such as Blauner (1964) and Edwards (1979) concerning developments in modes of workplace organization and their impacts on workers' experiences.



Checklist

Content analysis

-
- ☐ Have you clearly defined your research questions?
 - ☐ Is the population of documents to be content analysed relevant to your research questions?
 - ☐ Can you justify your sampling approach?
 - ☐ Have you made sure that your dimensions do not overlap?
 - ☐ Have you made sure that the categories used for each of your dimensions do not overlap?
 - ☐ Are the categories you use for each dimension exhaustive?
 - ☐ Do all the dimensions allow you to answer your research questions?
 - ☐ Have you piloted your coding schedule?
 - ☐ Are the coding instructions clear?
 - ☐ If your research is based on the mass media, can you justify the time span of your coverage?
 - ☐ Are you clear about the unit of analysis?



Key points

- Content analysis is very much located within the quantitative research tradition of emphasizing measurement and the specification of clear rules that exhibit reliability.
- While traditionally associated with the analysis of mass-media content, it is in fact a very flexible method that can be applied to a wide range of phenomena.
- It is crucial to be clear about your research questions in order to be certain about your units of analysis and what exactly is to be analysed.
- You also need to be clear about what is to be counted.
- The coding schedule and coding manual are crucial stages in the preparation for a content analysis.
- Content analysis becomes particularly controversial when it is used to seek out latent meaning and themes.



Questions for review

- To what kinds of documents and media can content analysis be applied?
- What is the difference between manifest and latent content? What are the implications of the distinction for content analysis?

What are the research questions?

- Why are precise research questions especially crucial in content analysis?
- With what general kinds of research questions is content analysis concerned?

Selecting a sample

- What special sampling issues does content analysis pose?

What is to be counted?

- What kinds of things might be counted in the course of doing a content analysis?
- To what extent do you need to infer latent content when you go beyond counting words?

Coding

- Why is coding so crucial in content analysis?
- What is the difference between a coding schedule and a coding manual?
- What potential pitfalls need to be guarded against when devising coding schedules and manuals?

Advantages of content analysis

- 'One of the most significant virtues of content analysis is its immense flexibility in that it can be applied to a wide variety of documents.' Discuss.

Disadvantages of content analysis

- To what extent does the need for coders to interpret meaning undermine content analysis?
- How far are content analysis studies atheoretical?